**The Product Recommendation for H&M**

Course: BIA 679

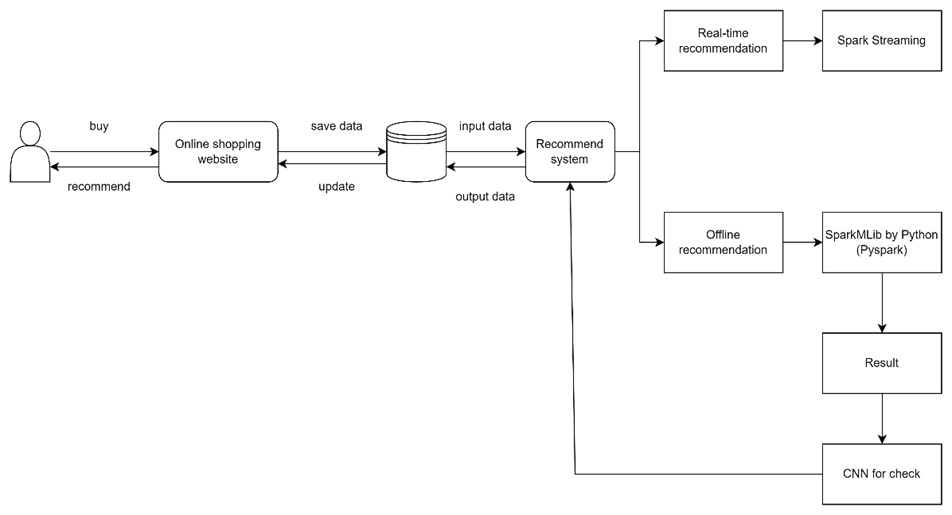
Member: Tengyue Chen, Hao Miao, Haoxing Zhang

**Introduction**

Online shopping has become one of the shopping channels for consumers with the development of the Internet. Total e-commerce sales in 2021 are expected to be $870.8 billion and grow 14.2% from 2020 (Young, 2022) in the U.S. Many online shopping companies have grown into well-known global companies, such as Amazon, Ebay, and Alibaba. However, with the increase in the number of items on online shopping platforms or websites. However, because the number of products on online shopping platforms continues to increase, the users may not choose their favorite products.

Recommend systems provide users with product information to help users make decisions and complete online purchases. The E-commerce recommendation system builds models that reflect user attributes and behaviors through the collected user information (Zhao, 2019). Online shopping platforms could use the E-commerce recommendation model in the backend to help users quickly find their favorite products.

This project attempts to build a model based on the dataset of products and user behavior provided by H&M. This model could predict the user's potential product selection and provide purchase suggestions through the user's previous purchase behavior or habits. Because real-time data cannot be obtained, the model of this project will focus on offline recommendation. The figure 1 shows the design of the recommendation model system.

****

Figue1. Model Design

**Data collection**

The dataset is collected from the Kaggle(<https://www.kaggle.com/competitions/h-and-m-personalized-fashion-recommendations/data>). The dataset contains the purchase history of H&M customers in the online store. The dataset contains three files, includes product pictures, product purchase records, and user information. Table 1 and table 2 show the dataset columns.

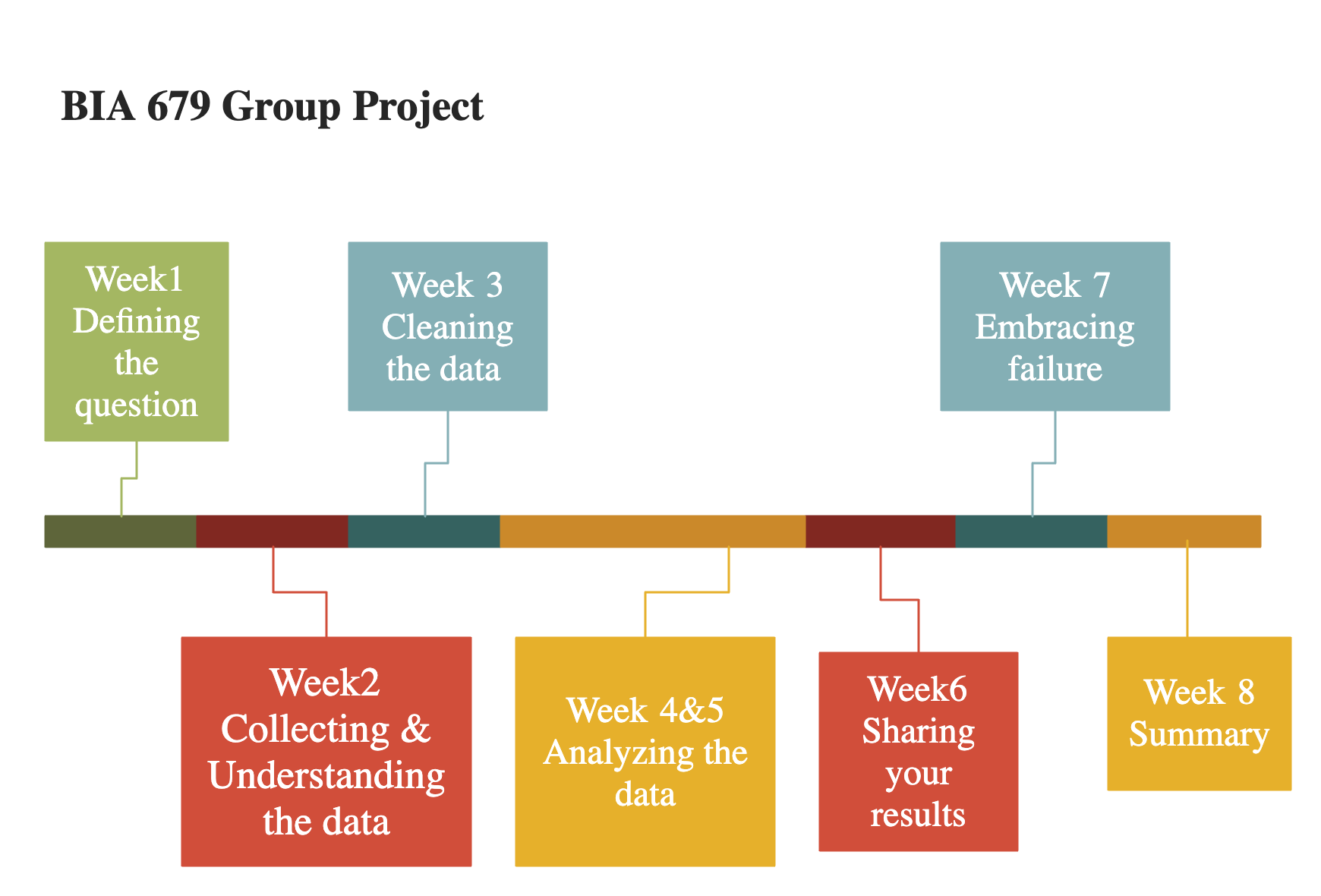
|  |  |
| --- | --- |
| Columns | Feature |
| article\_id | The article id |
| product\_code | The code of product |
| prod\_name | The name of product |
| product\_type\_no | The number of product type |
| product\_type\_name | The name of product name |
| product\_group\_name | The group name of product |
| graphical\_appearance\_no | The number of graphical appearances |
| graphical\_appearance\_name | The name of graphical appearances |
| colour\_group\_name | The name of color group |
| perceived\_colour\_value\_id | The value id for perceived color |
| perceived\_colour\_master\_id | The master id for perceived color |
| perceived\_colour\_master\_name | The master’s name for perceived color |
| department\_no | The number of departments |
| department\_name | The name of departments |
| index\_code | The index code |
| index\_name | The name of index |
| index\_group\_no | The number of index group |
| index\_group\_name | The name of index group |
| section\_no | The number of sections |
| section\_name | The name of section |
| garment\_group\_no | The number of garment group |
| garment\_group\_name | The name of garment group |
| detail\_desc | The detail describes |

Table 1. Columns of articles

|  |  |
| --- | --- |
| Columns | Feature |
| customer\_id | The customer id |
| FN |  |
| Active | Active or not active |
| club\_member\_status | The status of club member |
| fashion\_news\_frequency | The frequency of fashion news |
| age | The customer age |
| postal\_code | The customer’s postal code |

Table 2. Columns of customer

**Project Timeline**



**Teamwork Log**

Week 1 Teamwork Log

|  |  |  |  |
| --- | --- | --- | --- |
| Time | Event | Member | Duration |
| Monday | Slide Generate | All | 10min |
| Tuesday | Discussion of project | All | 1.5hr |
| Tuesday | Slides for project | Tengyue | 45min |
| Wednesday | Log generate | Haoxing, Hao | 15min |

Week 2 Teamwork Log

|  |  |  |  |
| --- | --- | --- | --- |
| Time | Event | Member | Duration |
| Tuesday | Discussion on group due | all | 30min |
| Tuesday | White Paper | All | 1hr |
| Tuesday | Slides | All | 30min |

Reference:

Zhao, Xuesong. “A Study on e-Commerce Recommender System Based on Big Data.” *2019 IEEE 4th International Conference on Cloud Computing and Big Data Analysis (ICCCBDA)*, 2019,<https://doi.org/10.1109/icccbda.2019.8725694>.

Jessica Young | Feb 18, 2022, et al. “US Ecommerce Grows 14.2% in 2021.” *Digital Commerce 360*, 16 Sept. 2022,<https://www.digitalcommerce360.com/article/us-ecommerce-sales/>.

“H&M Personalized Fashion Recommendations.” *Kaggle*, https://www.kaggle.com/competitions/h-and-m-personalized-fashion-recommendations/data.